

Hierarchical modelling of faecal egg counts to assess anthelmintic efficacy

Michaela Paul

Institute of Mathematics, University of Zurich, Zurich, Switzerland

michaela.paul@uzh.ch

Paul R. Torgerson

Section of Epidemiology, Vetsuisse Faculty, University of Zurich, Zurich, Switzerland

Johan Höglund

Department of Biomedical Sciences and Veterinary Public Health,
Swedish University of Agricultural Sciences, Section for Parasitology, Uppsala, Sweden

Reinhard Furrer

Institute of Mathematics, University of Zurich, Zurich, Switzerland.

January 14, 2014

Abstract

Counting the number of parasite eggs in faecal samples is a widely used diagnostic method to evaluate parasite burden. Typically a sub-sample of the diluted faeces is examined for eggs. The resulting egg counts are multiplied by a specific correction factor to estimate the mean parasite burden. To detect anthelmintic resistance, the mean parasite burden from treated and untreated animals are compared. However, this standard method has some limitations. In particular, the analysis of repeated samples may produce quite variable results as the sampling variability due to the counting technique is ignored. We propose a hierarchical model that takes this sampling variability as well as between-animal variation into account. Bayesian inference is done via Markov chain Monte Carlo sampling. The performance of the hierarchical model is illustrated by a re-analysis of faecal egg count data from a Swedish study assessing the anthelmintic resistance of nematode parasite in sheep. A simulation study shows that the hierarchical model provides better classification of anthelmintic resistance compared to the standard method.

Keywords: Anthelmintic resistance; Bayesian methods; Faecal egg count reduction test (FECRT); Markov chain Monte Carlo (MCMC); Sub-sampling.

1 Introduction

The extensive use of anthelmintic drugs to expel parasitic worms has lead to an increasing problem of anthelmintic resistance (AR) particularly to gastrointestinal nematodes in every livestock host (Kaplan, 2004). Reliable methods to assess and to detect changes in drug efficacy are thus essential for effective control of parasitic infections. Various in-vivo and in-vitro methods have been developed to detect AR (see review by Demeler et al., 2012). The most widely used (in-vivo) method is the faecal egg count reduction test (FECRT). It compares the number of parasite eggs in faecal samples taken before and after treatment (McKenna, 1990), or in samples of treated and untreated animals as control group (Coles et al., 1992, 2006), based on the assumption that the faecal egg counts (FECs) adequately reflect the

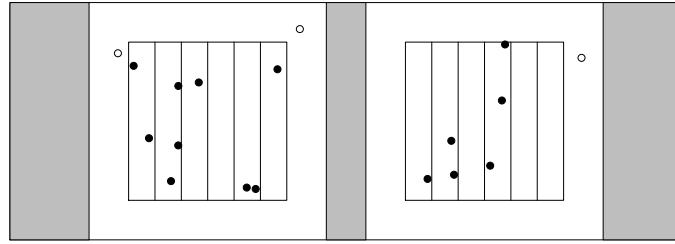


Figure 1: Schematic representation of a McMaster slide. Eggs lying within the two grid areas (filled circles) are counted to calculate the number of eggs per gram of faeces. With a correction factor of 50, this yields $(9 + 6) \times 50 = 750$ epg.

presence of adult worms in the gastrointestinal tract. For sheep and goats, resistance is declared if the percentage reduction in mean FECs is less than 95% and if the lower limit of a corresponding 95% confidence interval (CI) for the reduction in means is below 90%. If only one of the two criteria is met, resistance is suspected (Coles et al., 1992). For other livestock, different thresholds have been suggested.

A major advantage of the FECRT is its easy and straightforward use for all three available broad-spectrum anthelmintic groups. Standardised methods on how to obtain and compare FECs from various hosts have been advocated by the World Association for the Advancement of Veterinary Parasitology (WAAVP) (Coles et al., 1992). Typically, variants of the McMaster egg counting technique as detailed in Coles et al. (1992) are used to obtain FECs. First, the sample of faeces is diluted with a flotation fluid, thoroughly mixed and sieved to remove large debris. A sample of the resulting solution is then put in a counting chamber (see Figure 1). The eggs will float to the surface and can be counted. Subsequently, the number of eggs per gram of faeces (epg) is calculated by multiplying the number of eggs observed within the grid areas by an appropriate factor which depends on the amount of faeces and flotation fluid used, as well as the total volume of the McMaster chambers ($2 \times 0.15\text{ml}$). Most commonly, faeces are diluted by a factor of 15, i.e. the method has an analytic sensitivity or detection level of $15/0.3 = 50$ epg.

The detection level influences the accuracy of the FECRT. Results from repeated tests can be quite variable (Miller et al., 2006) particularly with low pre-treatment FECs, and are not reliable if only a small proportion ($\leq 25\%$) of the worms are resistant (Martin et al., 1989). A reason for this is that the egg counting technique introduces (substantial) variability which usually is not accounted for (Torgerson et al., 2012). Furthermore, the distribution of parasites and egg counts between animals is typically overdispersed (Grenfell et al., 1995). Despite the limitations of the standard FECRT, only few authors have proposed more elaborate statistical models for the analysis of overdispersed FECs. For instance, Torgerson et al. (2005) assumed a negative binomial distribution for FECs from treated and untreated animals and used parametric bootstrap to compute 95% CIs for the reduction in mean. Denwood et al. (2010) assumed a Poisson-gamma model for the FECs, where the post-treatment means were linked to the pre-treatment means via scale factors. Inference was done via Markov chain Monte Carlo (MCMC, e.g., Gilks et al., 1996) using JAGS software (Plummer, 2003).

To the best of our knowledge, no attempts have been made so far to also incorporate the variability of the egg counting process. In this paper, we propose a hierarchical model for the analysis of FECs that accounts for the above mentioned variabilities. The model is similar in spirit to econometric models for underreported counts which often occur in survey data due to varying reporting probabilities (e.g. Winkelmann, 2008; Fader and Hardie, 2000). Bayesian inference is done via tailored MCMC. In the following, we briefly review the conventional FECRT and summarise a recent study that was undertaken to get an updated status of the prevalence of AR of nematode parasites in Swedish sheep flocks (Höglund et al., 2009). We then discuss our hierarchical model formulation and computational approach for inference, followed by a re-analysis of the Swedish sheep FECs and a simulation study comparing the ability of the FECRT and the hierarchical model to detect AR. We close with some discussion.

2 Faecal egg count reduction test

The WAAVP guidelines (Coles et al., 1992) detail a McMaster egg counting method with analytical sensitivity of 50 epg. To somewhat lessen the effect of the counting technique on the FECRT, they suggest to use groups of (at least) 10 animals, where the individual pre-treatment eggs should be ≥ 150 . One group of n_T animals is treated, and the other group of n_C animals serves as control. The percentage reduction is then calculated as

$$\text{FECRT} = 100 \left(1 - \frac{\bar{x}_T}{\bar{x}_C} \right), \quad (1)$$

where \bar{x}_T and \bar{x}_C denote the mean counts in the treated and control group, respectively. Assuming independence, an approximate variance of the log ratio is given by

$$\text{Var} \left(\log \frac{\bar{x}_T}{\bar{x}_C} \right) = \frac{1}{\bar{x}_T^2} \text{Var}(\bar{x}_T) + \frac{1}{\bar{x}_C^2} \text{Var}(\bar{x}_C). \quad (2)$$

This variance is used to construct a corresponding approximate 95% CI on the log-scale using the 97.5% quantile of a t -distribution with $n_T + n_C - 2$ degrees of freedom, which is then backtransformed to give a 95% CI for the FECRT (1). For an anthelmintic to be fully effective, no worms should survive treatment and any viable eggs after treatment indicate that some worms may have been resistant. However, a 95% or larger reduction in egg counts indicates that the anthelmintic treatment is still beneficial (Coles et al., 2006). As a decision rule, the WAAVP guidelines suggest that AR be considered present in sheep and goats if a) the reduction after treatment is less than 95% and b) the lower 95% confidence limit is below 90%.

Various (minor) variants of the FECRT and decision rules have been suggested in the literature depending on the study design and host species, see e.g. Cabaret and Berrag (2004); Miller et al. (2006). For instance, without a distinct control group and faecal samples taken from the same animals before and after treatment, McKenna (1990) proposed to calculate egg count reduction as in (1) with \bar{x}_C now denoting pre-treatment counts. A corresponding 95% CI may be constructed using (nonparametric) bootstrap (Davison and Hinkley, 1997). In practice, CIs for this paired situation are also constructed using variance (2), although this is not optimal from a statistical point of view.

3 Anthelmintic resistance of nematode parasites in Swedish sheep flocks

To assess the prevalence of anthelmintic resistance in Swedish sheep, 45 randomly selected farms with a minimum of 20 ewes were visited throughout Sweden during the grazing season of 2006 and 2007 (Höglund et al., 2009). From each farm two flocks of approximately 15 lambs were dewormed with either benzimidazole (BZ) or macrocyclic lactone. In this paper, we only focus on the BZ treated flocks. Faecal samples were taken before treatment and analysed with a McMaster with analytic sensitivity of 50 epg. Out of the 45 flocks, 39 flocks with a mean of ≥ 50 epg were re-sampled 7–10 days after treatment and also analysed with a McMaster. Flock sizes varied between 10 and 17 lambs and pre-treatment mean eggs were often low (ranging from 50 to 5580, with a median of 324 epg).

Mean egg count reduction was calculated using (1). Both approximate 95% CIs according to the WAAVP guidelines as well as 95% bootstrap CIs were calculated to assess anthelmintic efficacy. In 35 flocks, post-treatment FECs were all zero resulting in a FECRT of 100%. Of the remaining four flocks (IDs 24, 33, 36, and 39), BZ resistance was declared for two flocks (33 and 39). Subsequent investigation of pooled pre- and post-treatment larval cultures showed that *Haemoncus contortus* was the main species involved in resistance. BZ resistance of *H. contortus* was further investigated with molecular tests of pre-treatment cultures. According to the molecular data, resistance was indicated by an estimated allele

frequency of $\geq 95\%$ in a total of five flocks (flocks 33 and 39 as already indicated by the FECRT, as well as flocks 24, 36, and 37). In summary, Höglund et al. (2009) concluded that the clinical resistance status of Swedish sheep nematodes is still relatively low though resistance may be more widespread than indicated by the conventional FECRT.

4 Modelling approach

Suppose we have a group of n untreated animals. A faecal sample from each animal i is analysed by a McMaster method with analytic sensitivity f . If the required amount of faeces is not available for certain animals, this results in larger correction factors f_i for those animals. For notational simplicity we assume throughout the remainder of the text that the the same correction factor is used for all samples. A fixed proportion $p = 1/f$ of the diluted faecal suspension is put on a McMaster slide and the eggs within the grid areas are counted. Denote those observed raw counts by Y_i^{*b} , $i = 1, \dots, n$, where the superscript b indicates that counts are collected before anthelmintic treatment. Given the true number of eggs per gram of faeces, Y_i^b , the observed number of eggs Y_i^{*b} is binomial distributed with probability p and size Y_i^b . The Y_i^b are assumed to be Poisson distributed with mean μ_i^b . Heterogeneity between animals is taken into account by letting μ_i^b follow a gamma distribution with shape parameter ϕ and rate parameter ϕ/μ having mean μ and variance μ^2/ϕ . This yields the following pre-treatment model

$$\begin{aligned} Y_i^{*b} | Y_i^b &\sim \text{Bin}(Y_i^b, p), \\ Y_i^b | \mu_i^b &\sim \text{Pois}(\mu_i^b), \\ \mu_i^b | \phi, \mu &\sim \text{Gamma}(\phi, \phi/\mu). \end{aligned} \quad (3)$$

If the treatment is effective, the number of eggs in faecal samples taken from the same animals some days after treatment should be vastly reduced. After treatment, the epg rate in (3) is reduced by a factor δ . This yields

$$\begin{aligned} Y_i^{*a} | Y_i^a &\sim \text{Bin}(Y_i^a, p), \\ Y_i^a | \mu_i^b, \delta &\sim \text{Pois}(\delta \mu_i^b), \end{aligned} \quad (4)$$

where superscripts a are now used to indicate the situation after anthelmintic treatment. If the treatment is completely ineffective the underlying true epg rates after treatment should not change, whereas they should largely decrease in the case of an effective treatment. We thus assign a beta prior to the reduction parameter, $\delta \sim \text{Beta}(a_\delta, b_\delta)$. Gamma priors are assigned to both the dispersion parameter and the population epg rate, i.e. $\phi \sim \text{Gamma}(a_\phi, b_\phi)$ and $\mu \sim \text{Gamma}(a_\mu, b_\mu)$.

Typical worm burden and FECs differ depending on the animals (e.g. sheep, cattle) and types of parasites considered. Here we discuss the hyperparameters that will be used in the application in Section 6. The Swedish study investigated AR in sheep nematodes, with the majority of eggs due to trichostrongylid infections. Several authors have investigated the abundance and distribution of trichostrongylid eggs in sheep faeces. For instance, Grenfell et al. (1995) observed that overdispersion is correlated to the magnitude of the FEC and approaches a plateau for high mean egg counts. This is in line with the results by Morgan et al. (2005) who investigated faeces of 14 groups of commercially farmed sheep. Mean FECs ranged from 43 to 1915, and the estimated overdispersion parameter of a negative binomial distribution ranged from 0.18 (95%-CI: 0.10–0.32) to 2.3 (95%-CI: 0.2–4.2).

We thus assume a $\text{Gamma}(1, 0.7)$ as prior distribution for the overdispersion parameter ϕ . Then, 90% of the prior probability mass lies between 0.1 and 4.3 with a prior median of 1. For the population epg rate μ a more dispersed $\text{Gamma}(1, 0.001)$ distribution is assumed, where 90% of the prior probability mass lies between 51 and 2996. Finally, we assign a $\text{Beta}(1, 1)$ prior to the reduction parameter δ , so that all values between 0 and 1 are equally likely a priori. All priors are assumed to be (conditionally) independent. The influence of the prior distribution for δ on the results is also investigated in a small sensitivity analysis.

5 Markov chain Monte Carlo inference

Full Bayesian inference is based on the joint posterior distribution which is proportional to

$$\begin{aligned} \prod_{i=1}^n & \left\{ \binom{y_i^b}{y_i^{*b}} p^{y_i^{*b}} (1-p)^{y_i^b - y_i^{*b}} \binom{y_i^a}{y_i^{*a}} p^{y_i^{*a}} (1-p)^{y_i^a - y_i^{*a}} \right. \\ & \times \frac{(\mu_i^b)^{y_i^b}}{y_i^b!} \exp(-\mu_i^b) \frac{(\delta \mu_i^b)^{y_i^a}}{y_i^a!} \exp(-\delta \mu_i^b) \frac{(\phi/\mu)^\phi}{\Gamma(\phi)} \mu_i^{b\phi-1} \exp\left(-\frac{\phi}{\mu} \mu_i^b\right) \Big\} \\ & \times \mu^{a_\mu-1} \exp(-b_\mu \mu) \phi^{a_\phi-1} \exp(-b_\phi \phi) \delta^{a_\delta-1} (1-\delta)^{b_\delta-1}. \end{aligned}$$

Posterior marginal distributions for the parameters of interest cannot be analytically derived, and we use MCMC sampling for inference. If the full conditional distribution for a specific parameter is proportional to a known distribution Gibbs sampling (Casella and George, 1992) is applied. Otherwise, samples from the full conditional distribution are obtained using a Metropolis-Hastings (MH) algorithm (Chib and Greenberg, 1995) with a suitably chosen proposal distribution.

We start with the parameters for which Gibbs sampling is possible. The full conditionals for the true number of eggs before treatment, y_i^{*b} , $i = 1, \dots, n$, are (displaced) Poisson distributions (Staff, 1964) with mean $(1-p)\mu_i^b$ for $y_i^b \geq y_i^{*b}$ and zero probability for $y_i^b = 0, 1, \dots, y_i^{*b} - 1$. Similarly, the full conditionals for the true number of eggs after treatment, y_i^{*a} , $i = 1, \dots, n$, are (displaced) Poisson distributions with mean $(1-p)\delta\mu_i^b$ for $y_i^a \geq y_i^{*a}$. To update the individual epg rates before treatment, μ_i^b , we draw samples from a gamma distribution with shape $y_i^b + y_i^a + \phi$ and rate $\sum_{i=1}^n \mu_i^b + b_\delta$.

The full conditionals for the remaining three parameters have no analytically closed form. For updating the overdispersion parameter ϕ we draw samples from the full conditional

$$p(\phi | \cdot) \propto \frac{\phi^{n\phi+a_\phi-1}}{\Gamma(\phi)^n} \mu^{-n\phi} \left(\prod_{i=1}^n \mu_i^b \right)^{\phi-1} \exp\left(-\phi(1/\mu \sum_{i=1}^n \mu_i^b + b_\phi)\right)$$

using a MH algorithm with uniform proposal distribution centred around the current value (denoted by ϕ^0) and suitably truncated to ensure that $\phi > 0$, i.e. proposal values are drawn from a $U(\max\{0, \phi^0 - s\}, \phi^0 + s)$ distribution (with tuning parameter s) and accepted with probability $\min\{1, p(\phi^* | \cdot)/p(\phi^0 | \cdot) \times q(\phi^0 | \phi^*)/q(\phi^* | \phi^0)\}$.

For the population epg rate μ we have

$$p(\mu | \cdot) \propto \mu^{-(n\phi-a_\mu+1)} \exp\left(-b_\mu \mu - \frac{\phi \sum_{i=1}^n \mu_i^b}{\mu}\right) = \exp\left(-a \log(\mu) - b_\mu \mu - \frac{c}{\mu}\right) \quad (5)$$

with $a = n\phi - a_\mu + 1$, $b = b_\mu$, and $c = \phi \sum_{i=1}^n \mu_i^b$. To sample from this full conditional distribution, we use a MH algorithm with a suitable approximation of (5) as proposal distribution. The full conditional may be approximated by an inverse gamma, log-normal or gamma distribution with respective parameters chosen to match the mode and the curvature at the mode of (5). Further details about these approximations are given in Appendix A.1.

To determine which approximation is best, we compute the Kullback-Leibler divergence (Kullback and Leibler, 1951) between the full conditional (5) and the approximating inverse gamma/log-normal/gamma distribution. In general, an inverse gamma distribution provides the best approximation if a is large (say, $a > 2$). For smaller a , a log-normal distribution is often suitable.

Finally, the full conditional for the reduction in mean δ is

$$p(\delta | \cdot) \propto \delta^{\sum_{i=1}^n y_i^a + a_\delta - 1} (1-\delta)^{b_\delta - 1} \exp\left(-\delta \sum_{i=1}^n \mu_i^b\right). \quad (6)$$

As proposal distribution in a MH algorithm, we use a beta distribution with parameters chosen to match the mode and curvature at the mode of the above full conditional distribution, see Appendix A.2.

The methods are implemented in an R package `eggCounts` available from <http://www.math.uzh.ch/as> or <http://cran.r-project.org/web/packages/eggCounts/>.

Table 1: FECR results for the five BZ treated flocks for which the molecular data indicated anthelmintic resistance. Shown are the results of the FECRT (based on Eq. (1), together with 95% approximate confidence intervals and bootstrap percentile intervals), as well as the results of the paired model (posterior median of $100(1 - \delta)$ and 95% HPD interval). Results are shaded in grey if resistance is declared, and printed in italics if resistance is suspected based on the WAAVP guidelines.

flock	FECRT			hierarchical model	
	FECR	approximate CI	bootstrap CI	median	HPD interval
24	99.0	[96.3, 99.8]	[97.5, 99.9]	99.0	[98.5, 99.4]
33	82.0	[65.3, 90.6]	[71.3, 89.4]	82.0	[77.7, 86.0]
36	97.5	[90.6, 99.4]	[93.9, 100.0]	97.1	[94.0, 99.3]
37	100.0			<i>97.7</i>	<i>[90.0, 100.0]</i>
39	92.3	[62.9, 98.4]	[83.8, 97.3]	92.0	[88.9, 94.8]

6 Application to Swedish FECRT study

In the following, the FECs of 39 sheep flocks by Höglund et al. (2009) are re-analysed with the hierarchical model discussed in Section 4. For 28 out of the 575 BZ treated animals no post-treatment FEC is available. Most of those animals had a pre-treatment FEC of zero or one. For the analysis, all animals with missing post-treatment FECs were excluded. In addition, one animal had a pre-treatment epg of 30 which is not possible with a correction factor of 50. Here, 3 eggs floating outside the grid areas of the McMaster slide were counted with a factor of 10. As only the eggs within the grid areas should be counted, we set this FEC to zero.

For each flock, we obtained 10 000 posterior samples using a burn-in period of 10 000 iterations and a thinning of 10. The tuning parameter for the update of ϕ was selected to achieve an acceptance rate between 30 and 40 per cent. Acceptance rates for μ and δ were generally high with an average of 98.8% (range: 92.8–100.0%) and 98.1% (range: 97.4–100.0%). Standard output diagnostics provided by the R-package *coda* (Plummer et al., 2006) were applied to check for convergence.

Table 1 shows the result of the FECRT as discussed in Section 2 for the five BZ treated flocks for which the molecular data indicated anthelmintic resistance. In addition to the approximate 95% confidence intervals based on (2) we have also used a nonparametric bootstrap approach with paired sampling implemented in the R package *boot* (Canty and Ripley, 2012) to compute 95% bootstrap percentile intervals based on 1999 bootstrap samples. As the former approach ignores that counts are taken from the same animals before and after treatment, the resulting approximate confidence intervals are wider than the respective bootstrap intervals. The estimated percentage reductions in mean epg rate ($100(1 - \delta)$) obtained with the paired model are quite similar to the standard FECRT. There is a clear indication of AR resistance for flocks 33 and 39. Post-treatment FECs for flock 37 were all zero, and thus the percentage reduction in mean epg rate is estimated as 100% by the FECRT, and no confidence intervals can be computed. As the pre-treatment mean epg was rather low for this flock, the sampling variability due to the McMaster method plays an important role and leads to a relatively wide HPD interval for the reduction in mean.

Figure 2 shows the estimated percentage reduction in mean epg rate together with 95% HPD intervals for all flocks, including the 35 flocks with zero post-treatment mean epgs for which the standard FECRT only provides an estimated reduction in mean of 100% without confidence intervals. The width of the intervals is largest for flocks 35 and 22 which had the lowest pre-treatment means (below 100 epg).

An advantage of the Bayesian inference approach is that we obtain not only point estimates (and confidence intervals), but also posterior marginal distributions for the parameters of interest. For instance, we can make probability statements about the true reduction in mean being below or above some

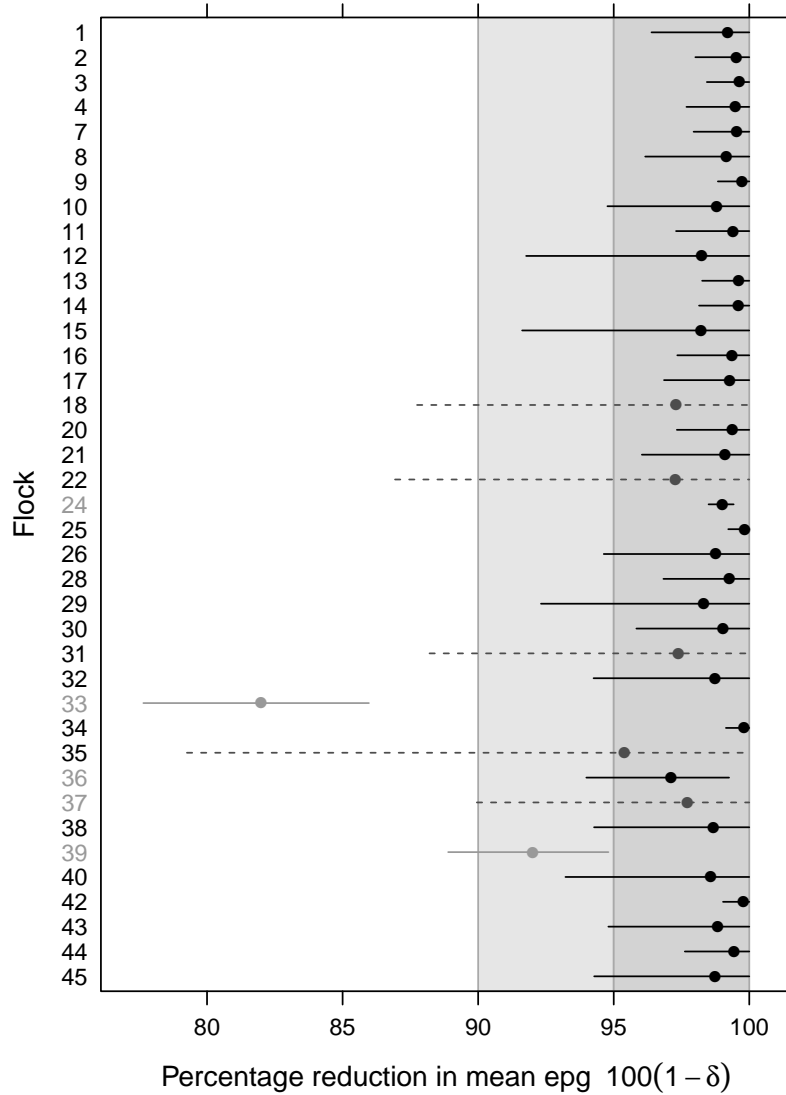


Figure 2: Estimated reduction in mean epg rate per flock treated with benzimidazole, together with 95% HPD intervals. Intervals are shown in grey if resistance is declared, and with dashed lines if resistance is suspected based on the WAAVP guidelines. The flock number on the y-axis is printed in grey if anthelmintic resistance is suspected based on the molecular testing.

threshold. These probabilities could then be used to judge AR (see Denwood et al., 2010). Figure 3 shows the posterior marginals for the reduction in mean epgs $1 - \delta$ for flocks 37 and 39. According to the WAAVP guidelines for nematodes in sheep, AR is present if the FECR is less than 95%. The posterior probability that the reduction in mean is less than 95% is shaded in grey in Figure 3. For flock 37, this probability equals 0.22 indicating that there might be some resistance present. In contrast, a probability of 0.99 for flock 39 provides strong evidence for AR.

We also investigated the sensitivity of the results to the prior assumptions. In total, four different prior assumptions for δ were considered:

1. an uninformative Beta(1, 1) distribution,
2. an informative Beta(0.5, 1) distribution with more weight on reductions close to 0 (i.e. no AR),

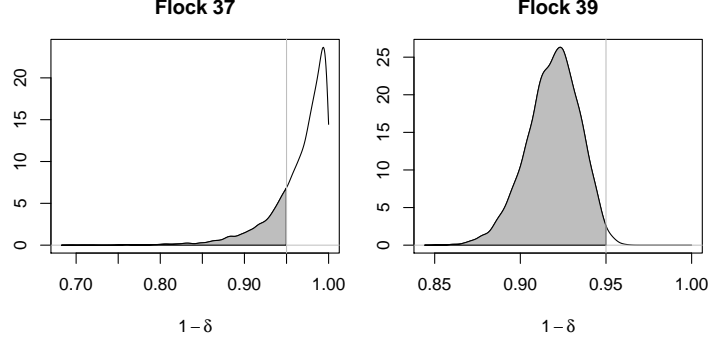


Figure 3: Posterior marginal distributions for the reduction in mean $(1 - \delta)$ for flocks 37 and 39. The probability that the reduction in mean eggs is less than 95% is shaded in grey.

3. an informative $\text{Beta}(1, 0.5)$ distribution with more weight on reductions close to 1,
4. a very informative $\text{Beta}(5, 1)$ distribution with 99% of the probability mass on values $\delta > 0.4$ (i.e. strong prior belief that AR is present).

The remaining model specifications stayed the same as before.

Figure 4 shows the four prior distributions (top left) and the resulting posterior marginal distributions for δ for the five BZ treated flocks for which the molecular data indicated AR. In general, the posteriors are very similar for the first three priors, and shifted somewhat to the right for the very informative prior. In the top of the plots the respective 95% HPD intervals are displayed. Using the HPD intervals to decide on AR based on the WAAVP criteria mostly leads to the same conclusions. Results are most sensitive to the prior assumptions for flock 37 with very low pre-treatment mean FECs and zero post-treatment FECs.

7 Simulation study

To further investigate the ability of the hierarchical model to assess anthelmintic efficacy compared to the standard FECRT, a simulation study was conducted. We first simulated a number of pre- and post-treatment FECs for different anthelmintic efficacies and analysed them with both the FECRT and hierarchical model. We then computed the percentage of how often each method declared presence or absence of AR according to the WAAVP guidelines.

7.1 Setup

Similar to the Swedish FECRT study we consider flocks of $n = 15$ animals. Pre-treatment FECs for these animals are simulated as follows: We first draw mean egg rates μ_i from a gamma distribution with shape 0.9 and rate 0.9/500 and simulate the number of eggs y_i^b from a Poisson distribution with this mean. So marginally, the number of eggs are drawn from a negative binomial distribution with mean 500 and overdispersion parameter 0.9. We then obtain the pre-treatment FECs (counted on a McMaster slide) by drawing from a Poisson distribution with mean y_i^b/f , where $f = 50$ corresponds to the correction factor of the McMaster method. Assuming an anthelmintic efficacy of $d\%$, the number of eggs after treatment y_i^a are simulated from a Poisson distribution with mean $\mu_i(1 - d/100)$. Post-treatment FECs are again obtained by sampling from a Poisson distribution with scaled mean y_i^a/f .

In total we considered 8 different scenarios with anthelmintic efficacies d ranging from 85 to 99%. For each scenario, 2000 data sets were simulated and analysed with the FECRT using a) an approximate CI ignoring the paired structure of the data, b) a paired bootstrap percentile interval, and c) the paired

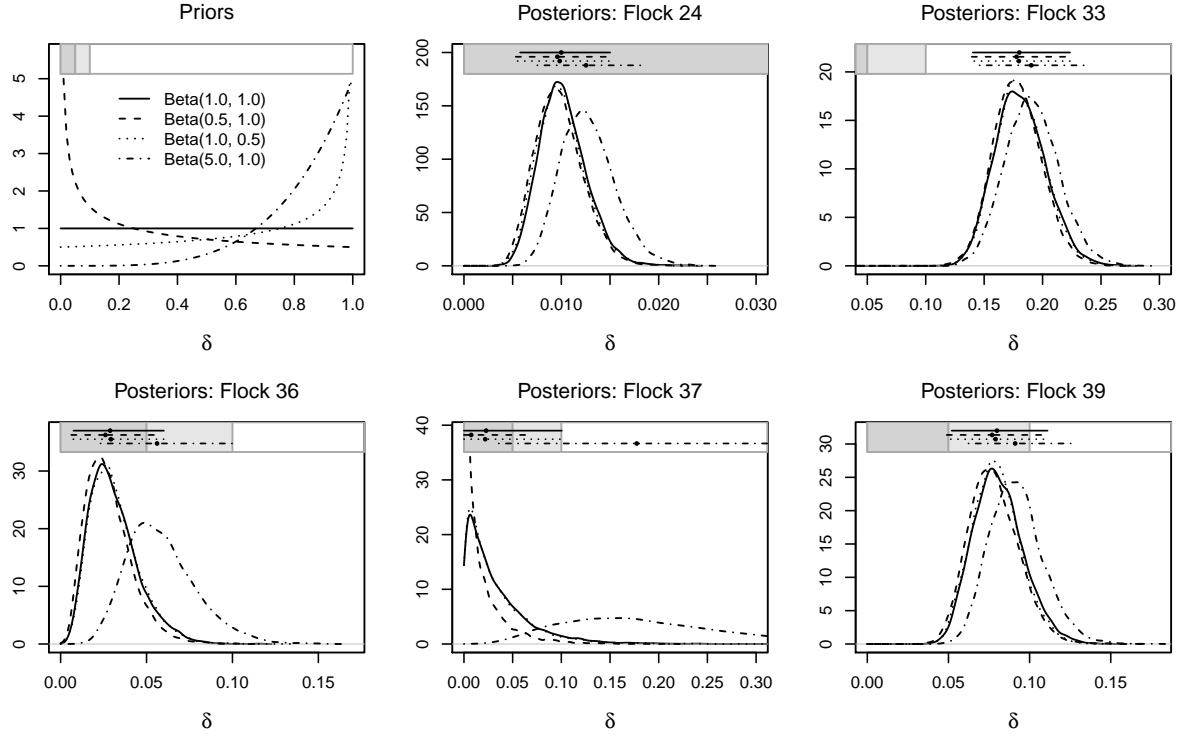


Figure 4: Sensitivity to the prior assumptions for reduction parameter δ for the five BZ treated flocks in Table 1. The top left plot shows four different prior distributions for δ . The remaining plots show the resulting posterior marginal distributions together with respective 95% HPD intervals in the top. In addition, the definition of declared and suspected AR based on the WAAVP guidelines is illustrated by the areas shaded in dark and light grey.

model (3)–(4). For the latter, the same priors as discussed in Section 4 for the Swedish FECR study were used. As before, we obtained 10 000 posterior samples using a burn-in period of 10 000 iterations and a thinning of 10. The tuning parameter for the update of ϕ was selected to achieve acceptance rates between 30 and 40 per cent.

7.2 Results

The left plot in Figure 5 shows how often the considered methods declared AR to be present based on the WAAVP guidelines, i.e. both the estimated reduction is less than 95% and the lower limit of a 95% confidence interval is less than 90%. Similarly the middle and right plot show how often AR is declared to be possible (one of the two criteria is met) and absent (none of the criteria is met). For a specific anthelmintic efficacy the percentages in the three plots sum to 1 for each method. For the FECRT methods, AR is considered to be absent if the estimated reduction equals 100% and no confidence interval can be computed due to a post-treatment variance of the FECs of 0. Suppose we consider the treatment as effective if the anthelmintic efficacy d is larger than 95%. This is indicated with a solid grey vertical line in Figure 5.

For large values of $d = 99\%$ and for low efficacies ($d = 85\text{--}87\%$), all methods yield the correct decision for nearly all simulated data sets. Closer to assumed efficacies of $d = 95\%$, the decision gets more difficult. In general, the hierarchical model performs better than the FECRT with bootstrap CIs and selects the true category of present (absent) AR when anthelmintic efficacies are smaller (larger) than 95% with highest percentage. However, the percentages of falsely deciding on absent AR when $d = 91$ or 93% are

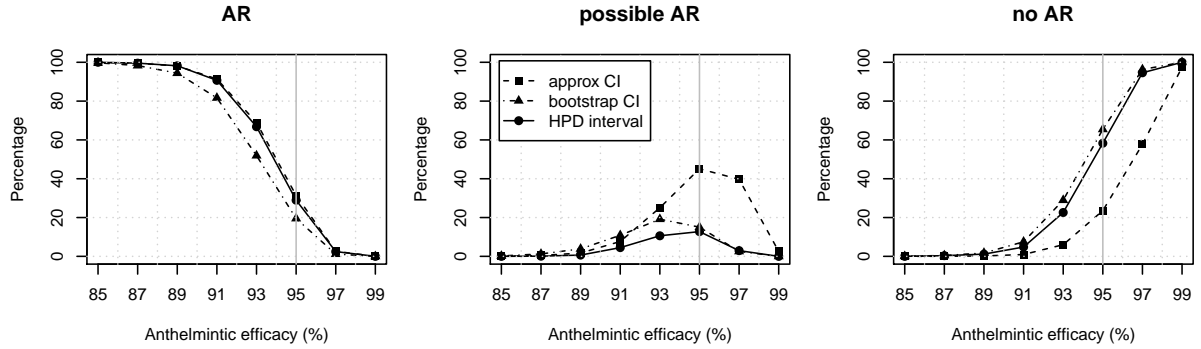


Figure 5: Results of the simulation study. Shown are percentages of how often the FECRT using an approximate CI (■) or bootstrap percentile intervals (▲) and model (4) (●) declared anthelmintic resistance to be present (both criteria of the WAAVP guidelines are met), possible (one of the criteria is met), or absent (none of the criteria is met) for true anthelmintic efficacies ranging from 85 to 99%.

lowest for the standard FECRT with approximate CIs. This is mainly due to the large width of the approximate CIs resulting in high percentages for the middle category where a clear decision on absence or presence of AR is not possible.

With the hierarchical model, an even better detection of AR can be obtained by using the posterior marginal distribution of δ for classification. Figure 6 shows posterior marginal probabilities that the reduction in mean epg rates ($1 - \delta$) is less than 95%. The same curve as for the HPD interval in Figure 5 can be seen. The further away from $d = 95\%$, the more the probabilities are concentrated towards 0 or 1. Denwood et al. (2010) suggested that criteria $\text{Prob}(1 - \delta < 0.95) > 0.975$ and $\text{Prob}(1 - \delta < 0.95) < 0.025$ could be used to decide on “confirmed resistance” and “confirmed susceptibility”. With such a classification rule, no false decisions on absent AR occur with efficacies $d < 91\%$, and only 1 and 10 false decisions (out of 2000) occur when $d = 91$ and 93% , respectively.

8 Discussion

Anthelmintic resistance of nematodes in livestock is becoming more widespread. The problem is serious in e.g. sheep nematodes and emerging e.g. in cattle nematodes. Adequate methods to detect and monitor resistance are thus needed. Most commonly FECs are used as a diagnostic tool to quantify parasite burden. While many attempts have been made to improve the diagnostic procedures for various hosts and parasites, the development of statistical methods for the analysis of FECs has received less attention. In this paper we have proposed a hierarchical model that is able to incorporate sources of variability inherent in FECs. In particular the variability resulting from the McMaster counting technique is taken into account when estimating anthelmintic efficacy. This is especially important in situations with low pre-treatment FECs and high correction factors. Then the standard FECRT quite often results in an estimated reduction in mean epg rate of 100% without confidence interval. In contrast, the hierarchical model provides a quantification of the uncertainty concerning the reduction parameter via the posterior marginal distribution.

For instance, in the application to the Swedish FECRT study anthelmintic resistance could not be ruled out for flock 37 based on the results of the hierarchical model. For this flock the molecular test also indicated AR. Note that the larval differentiation undertaken in this study provides estimates of the abundance of the different nematode species. In some flocks, *H. contortus* was the prevailing species whereas in other flocks *H. contortus* was only marginally present at the time of treatment with BZ. If

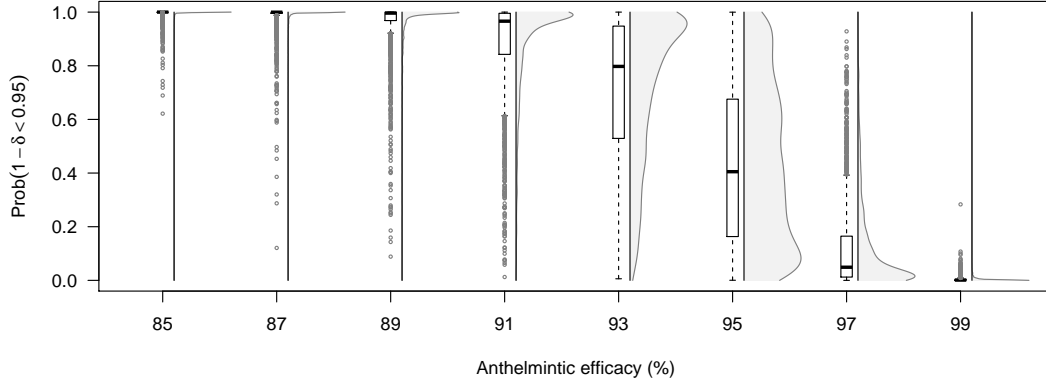


Figure 6: Results of the simulation study. Shown are boxplots and kernel density estimates of the posterior marginal probabilities that the reduction in mean eggs ($1 - \delta$) is less than 95% (indicating resistance) for true anthelmintic efficacies ranging from 85 to 99%.

only a small proportion of eggs stems from the resistant species, then it is very difficult to detect AR based on FECs.

An advantage of the hierarchical model formulation is its flexibility in model specification. In this paper we have assumed that the efficacy of the anthelmintic treatment is the same for each animal. This seems sensible as one would expect similar low efficacy in a resistant community. In a susceptible community efficacy should be high for all animals even though efficacy may be lower for some individuals due to poor metabolism or availability of the drug in these animals (Cabaret and Berrag, 2004). One could let the reduction parameter δ vary across animals but this would require adequate prior information about the efficacy distribution.

A further generalization of the model not considered here is the incorporation of zero-inflation. The WAAVP guidelines for sheep suggest to consider only animals with pre-treatment egg counts > 150 epg (Coles et al., 2006). Hence, animals with zero pre-treatment FECs are sometimes discarded for the analysis. However, those counts do provide information about worm burden and should not be neglected. For instance, to address excess zeros (from uninfected animals) we can replace the gamma distribution for the individual epg rates by a mixture of a gamma distribution and a point mass at zero.

Acknowledgements

The work of MP was supported by a grant from the Swiss National Science Foundation (ref: CR3313-132482). PT also recieved funding from the EU – KBBE.2011.1.3-04 288975 “Gloworm”. RF and MP acknowledge funding from the University Research Priority Program.

A Proposal distributions

A.1 Pre-treatment population mean μ

The unnormalized full conditional distribution of the pre-treatment population mean epg rate μ is given in (5) as $p(\mu|\cdot) \propto \exp(-G(\mu))$ with $G(\mu) = a \log(\mu) + b\mu + c/\mu$. The distribution has a mode at

$$m = \frac{-a + \sqrt{a^2 + 4bc}}{2b}, \quad (7)$$

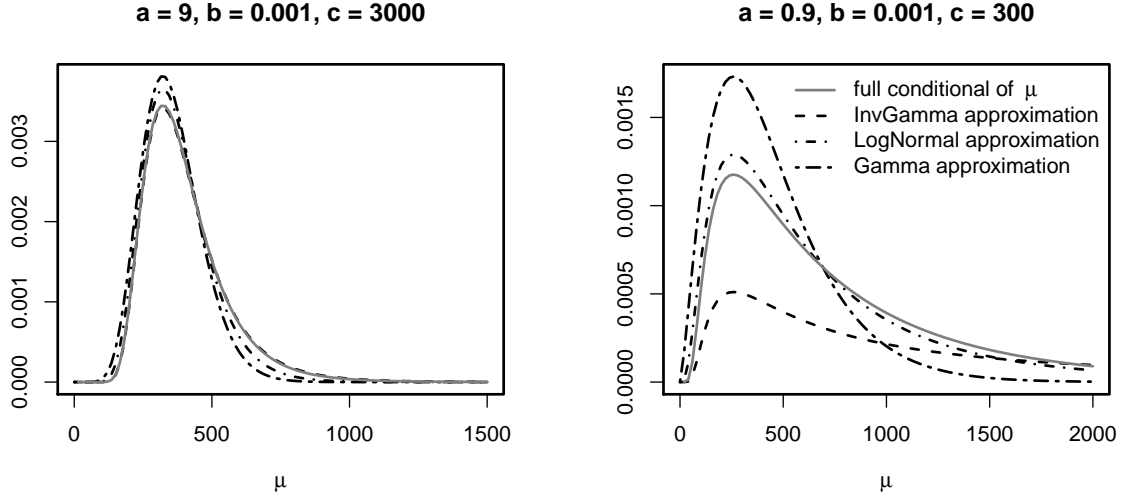


Figure 7: Gamma, inverse gamma, and log-normal approximation of the full conditional distribution of μ for two sets of parameters.

and the second derivative of minus the log density with respect to μ and evaluated at m is given by

$$G''(m) = -\frac{a}{m^2} + \frac{2c}{m^3}. \quad (8)$$

We can approximate the full conditional by matching the mode (7) and the second derivative (8) to the respective values of a known distribution.

For instance, an inverse gamma distribution with shape parameter α and scale parameter β has mode $m_{\text{IG}} = \beta/(\alpha + 1)$, and second derivative of minus the log density at the mode equal to $D_{\text{IG}} = -(\alpha + 1)/m_{\text{IG}}^2 + 2\beta/m_{\text{IG}}^3$. Replacing m_{IG} and D_{IG} by (7)–(8), and solving the two equations for the shape and scale parameter yields

$$\alpha = \beta/m - 1, \quad \beta = G''(m)m^3$$

as parameters of an inverse gamma proposal distribution for μ .

Analogously, we can approximate the full conditional distribution (5) by a log-normal distribution with mean $\log(m) + 1/G''(m)m^2$ and standard deviation $(\sqrt{G''(m)m})^{-1}$, or by a gamma distribution with shape $m\beta + 1$ and rate $G''(m)m$.

Figure 7 shows the full conditional together with the approximations for two different sets of parameter values. For the first parameter set, the inverse gamma distribution provides a very good approximation, whereas for the second set a log-normal distribution is better suited.

A.2 Reduction in mean δ

The unnormalised full conditional distribution of δ is given in (6) as

$$p(\delta | \cdot) \propto \exp\left(-[a \log(\delta) - b \log(1 - \delta) + c\delta]\right) = \exp(-G(\delta))$$

with $a = \sum_{i=1}^n Y_i^a + a_\delta - 1$, $b = b_\delta - 1$, and $c = \sum_{i=1}^n \mu_i^b$. The distribution has a mode at

$$m = -\frac{1}{2c}(-a - b - c + \sqrt{a^2 + 2ab - 2ca + b^2 + 2bc + c^2}) \quad (9)$$

and the second derivative of minus the log density at the mode m is given by

$$G''(m) = \frac{a}{m^2} + \frac{b}{(1-m)^2}. \quad (10)$$

Matching mode (9) and second derivative (10) to the respective values of a $\text{Beta}(\alpha, \beta)$ distribution yields

$$\alpha = (1 - m)m^2 G''(m) + 1, \quad \beta = m(m - 1)^2 G''(m) + 1$$

as parameters of an approximating beta distribution as proposal distribution for δ .

References

- Cabaret, J. and B. Berrag (2004). Faecal egg count reduction test for assessing anthelmintic efficacy: average versus individually based estimations. *Veterinary Parasitology* 121(1), 105–113.
- Canty, A. and B. Ripley (2012). boot: Bootstrap R (S-Plus) functions. R package version 1.3-4.
- Casella, G. and E. I. George (1992). Explaining the Gibbs sampler. *American Statistician* 46, 167–174.
- Chib, S. and E. Greenberg (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician* 49, 327–335.
- Coles, G. C., C. Bauer, F. H. M. Borgsteede, S. Geerts, T. R. Klei, M. A. Taylor, and P. J. Waller (1992). World Association for the Advancement of Veterinary Parasitology (WAAVP) methods for the detection of anthelmintic resistance in nematodes of veterinary importance. *Veterinary Parasitology* 44(1-2), 35–44.
- Coles, G. C., F. Jackson, W. E. Pomroy, R. K. Prichard, G. von Samson-Himmelstjerna, A. Silvestre, M. Taylor, and J. Vercruysse (2006). The detection of anthelmintic resistance in nematodes of veterinary importance. *Veterinary Parasitology* 136(3), 167–185.
- Davison, A. C. and D. V. Hinkley (1997). *Bootstrap Methods and Their Applications*. Cambridge: Cambridge University Press.
- Demeler, J., E. Schein, and G. von Samson-Himmelstjerna (2012). Advances in laboratory diagnosis of parasitic infections of sheep. *Veterinary Parasitology* 189(1), 52–64.
- Denwood, M. J., S. W. J. Reid, S. Love, M. K. Nielsen, L. Matthews, I. J. McKendrick, and G. T. Innocent (2010). Comparison of three alternative methods for analysis of equine Faecal Egg Count Reduction Test data. *Preventive Veterinary Medicine* 93(4), 316–323.
- Fader, P. S. and B. G. S. Hardie (2000). A note on modelling underreported Poisson counts. *Journal of Applied Statistics* 27(8), 953–964.
- Gilks, W. R., S. Richardson, and D. J. Spiegelhalter (1996). *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall.
- Grenfell, B. T., K. Wilson, V. S. Isham, H. E. G. Boyd, and K. Dietz (1995). Modelling patterns of parasite aggregation in natural populations: trichostrongylid nematode–ruminant interactions as a case study. *Parasitology* 111(S1), S135–S151.
- Höglund, J., K. Gustafsson, B.-L. Ljungström, A. Engström, A. Donnan, and P. Skuce (2009). Anthelmintic resistance in Swedish sheep flocks based on a comparison of the results from the faecal egg count reduction test and resistant allele frequencies of the β -tubulin gene. *Veterinary Parasitology* 161(1-2), 60–68.
- Kaplan, R. M. (2004). Drug resistance in nematodes of veterinary importance: a status report. *Trends in Parasitology* 20(10), 477–481.

- Kullback, S. and R. A. Leibler (1951). On information and sufficiency. *The Annals of Mathematical Statistics* 22(1), 79–86.
- Martin, P. J., N. Anderson, and R. G. Jarrett (1989). Detecting benzimidazole resistance with faecal egg count reduction tests and in vitro assays. *Australian Veterinary Journal* 66(8), 236–240.
- McKenna, P. B. (1990). The detection of anthelmintic resistance by the faecal egg count reduction test: An examination of some of the factors affecting performance and interpretation. *New Zealand Veterinary Journal* 38(4), 142–147.
- Miller, C. M., T. S. Waghorn, D. M. Leathwick, and M. L. Gilmour (2006). How repeatable is a faecal egg count reduction test? *New Zealand Veterinary Journal* 54(6), 323–328.
- Morgan, E. R., L. Cavill, G. E. Curry, R. M. Wood, and E. S. E. Mitchell (2005). Effects of aggregation and sample size on composite faecal egg counts in sheep. *Veterinary Parasitology* 131(1-2), 79–87.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*. March, pp. 20–22.
- Plummer, M., N. Best, K. Cowles, and K. Vines (2006). Coda: Convergence diagnosis and output analysis for MCMC. *R News* 6(1), 7–11.
- Staff, P. J. (1964). The displaced Poisson distribution. *Australian Journal of Statistics* 6(1), 12–20.
- Torgerson, P. R., M. Paul, and F. I. Lewis (2012). The contribution of simple random sampling to observed variations in faecal egg counts. *Veterinary Parasitology* 188, 397–401.
- Torgerson, P. R., M. Schnyder, and H. Hertzberg (2005). Detection of anthelmintic resistance: a comparison of mathematical techniques. *Veterinary Parasitology* 128(3-4), 291–298.
- Winkelmann, R. (2008). *Econometric Analysis of Count Data*. Berlin: Springer.